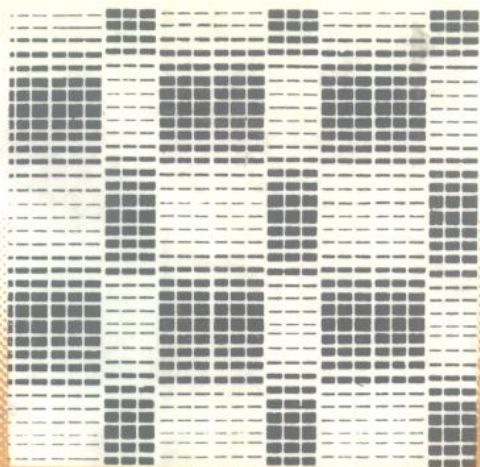




信息处理丛书·信息处理丛书

图像理解

王润生 编著

国防科技大学出版社

• 信息处理丛书 •

图 像 理 解

王润生 编著

国防科技大学出版社

DT11/2
内 容 简 介

本书重点介绍了图像理解的主要内容:线的提取、区域分割、纹理分析、形状描述、图像序列分析、双目图像分析和图像理解系统等。本书的特点是:按建立图像理解系统的主线安排章节,编入了国际上较新的研究成果,并融入了作者自己的一些工作成果,章末有小结并提供了进一步发展的见解。

本书可以作为信号与信息处理、计算机科学与工程、自动化、遥感等有关专业高年级学生和研究生的教学用书,也可供有关领域的科技工作者参考。

· 信息处理丛书 ·

图 像 理 解

王润生 编著

责任编辑:胡见堂

责任校对:张建军

*

国防科技大学出版社出版发行

邮政编码:410073 电话:(0731)4436564

新华书店总店科技发行所经销

湖南大学印刷厂印装

开本:850×1168 1/32 印张:12.8125 字数:321千

1995年10月第1版第1次印刷 印数1—3000册

ISBN 7-81024-313-6

TP·59 定价:16.80元

(本书如有印装质量问题,请与承印厂家直接联系)

8200178

前 言

图像理解是在图像处理基础上、结合人工智能与专家系统而发展起来的一门新兴学科。由于它可使智能机器增添视觉功能,在工业与国防建设等诸方面有着巨大的应用前景,所以一直为学术界和工业界所关注。有关的会议、杂志逐渐增加,新概念、新思想、新方法和系统都在陆续出现,也有了若干值得重视的应用。为了使更多的人了解这个学科的基本理论和研究方法,跟踪国际的学术发展,作者编写了这本书,以期为推动和繁荣我国的图像理解研究尽点力量。本书是在我 1987 年编写的《图像理解》讲义基础上,结合七期研究生教学实践写成的。由于本学科涉及的面太宽,资料浩繁,本人才疏学浅,书中难免存在不当甚至错误之处,恳请专家与读者给予指正。

当完成此书的时候,我首先要感谢美国麻省大学 A. Hanson 和 E. Riseman 两位教授,他们在我两次访美期间给了我许多启迪性的指导和热情的支持,使我对图像理解有了一个较深的了解。我还要感谢孙仲康、沈振康两位教授,是他们给我提供了首次讲解本书内容的机会。感谢校、系领导郭桂蓉教授、孙茂印教授的亲切关怀和支持,感谢梁甸农、沈振康两位教授审阅了本书。此外,我们研究小组的同志与同学都给了我不少帮助,在此一并表示谢意。

如果此书能给有志图像理解的读者一些帮助,本人将不胜欣慰。

作者谨识

1994. 8

目 录

第一章 绪 论

1.1	概述	1
1.1.1	什么是图像理解	1
1.1.2	图像理解的研究和应用	7
1.2	常用的数据结构	11
1.2.1	线性表	12
1.2.2	层次结构	13
1.2.3	图结构	16
1.2.4	复杂的循环结构举例	18
1.3	知识的表示与应用	20
1.3.1	常用的几种知识表示方法	21
1.3.2	不精确推理概述	23
1.4	本书章节安排	29

第二章 线

2.1	概述	31
2.2	简单的边缘检测算子	32
2.2.1	常见的检测局部边缘算子	33
2.2.2	简单边缘检测算子的性能比较	35
2.2.3	改进简单边缘检测算子的方法	39
2.3	Marr-Hildreth 边缘检测理论	47
2.3.1	基本原理	47
2.3.2	具体实现	52
2.3.3	多信道的综合	55
2.4	用 Facet 模型检测边缘	57
2.4.1	Facet 模型的形式化	58
2.4.2	利用 Facet 模型确定边缘	63
2.5	形成有意义的线特征	67
2.5.1	启发式连接	69
2.5.2	Hough 变换	73
2.5.3	相位编组方法	78

2.5.4	层次记号编组	80
2.6	小结	86
第三章 区 域		
3.1	概述	88
3.2	简单图像基于像元灰度的全局门限法	90
3.2.1	简单的直方图分析法	91
3.2.2	直方图凹形分析法	92
3.2.3	类别方差自动门限法	93
3.2.4	最佳熵自动门限法	95
3.2.5	矩不变自动门限法	98
3.2.6	极小误差自动门限法	100
3.3	简单图像基于像元邻域属性的全局门限法	102
3.3.1	基于边缘特性的方法	102
3.3.2	基于二阶统计特性的方法	107
3.3.3	利用两个“累积测量”选门限	110
3.3.4	松弛迭代方法	111
3.4	全局门限化直方图分割方法的评估	113
3.5	复杂图像基于图像属性的分割方法	118
3.5.1	自动多门限分割技术	118
3.5.2	循环分解法	123
3.5.3	分块分割再合并方法	124
3.6	区域连接增长	127
3.6.1	单连接区域增长	127
3.6.2	混合连接区域增长	128
3.6.3	中心连接区域增长	130
3.6.4	混合连接组合技术	132
3.7	分开—合并的分析方法	132
3.7.1	基本思想	132
3.7.2	处理过程	134
3.7.3	一致性谓词	135
3.8	彩色图像分割	137
3.9	小结	141
第四章 纹 理		
4.1	概述	145
4.2	空间灰度层共现矩阵	148

4.2.1	基本原理	148
4.2.2	共现矩阵应用举例	151
4.3	纹理能量测量	154
4.4	用 Markov 随机场模型描述纹理	158
4.5	用分形和分维理论描述纹理	165
4.5.1	分形和分维的基本概念	166
4.5.2	分形布朗模型用于图像分析	169
4.5.3	双毯求表面面积确定分形参量	174
4.6	纹理的结构分析方法和纹理梯度	178
4.6.1	纹理的结构分析方法	178
4.6.2	纹理梯度	180
4.7	小结	182
第五章 形 状		
5.1	概述	184
5.2	二维形状描述技术	186
5.2.1	内标量方法	187
5.2.2	外标量变换方法	192
5.2.3	内空间域技术	198
5.2.4	外空间域技术	202
5.2.5	二维形状描述方法回顾	203
5.2.6	二维形状的层次描述	204
5.3	三维物体的表示方法	207
5.3.1	骨架表示法	207
5.3.2	表面表示法	208
5.3.3	体积表示法和广义圆柱体	213
5.4	由图像性质导出表面方向	220
5.4.1	坐标系统和梯度空间	221
5.4.2	由正交投影图像获取三维形状信息	223
5.4.3	由体视投影图像获取三维形状信息	229
5.5	小结	234
第六章 图像序列分析		
6.1	概述	235
6.2	从图像序列中检测运动信息	236
6.2.1	简单的检测灰度变化的方法	236
6.2.2	基于灰度梯度确定瞬时速度场	239

6.3	基本记号的检测方法	243
6.3.1	点记号检测	243
6.3.2	点匹配	249
6.3.3	运动边缘检测	250
6.3.4	连通的像元集(小区域)的检测	252
6.4	流场几何	257
6.5	纯平移运动	262
6.5.1	基本关系	262
6.5.2	直接搜索平移轴	264
6.5.3	复对数变换分析方法	267
6.6	利用点对应性估计运动参数和三维结构	270
6.6.1	对曲面运动的估计	270
6.6.2	对平面运动的估计	275
6.7	利用局部光流场的分析方法	277
6.7.1	导出动力学关系	278
6.7.2	实现方法的考虑	281
6.8	全局光流技术	282
6.8.1	流场分割	283
6.8.2	估计三维运动和结构	286
6.9	小结	290

第七章 双目图像分析

7.1	概述	292
7.2	搜索对应性的条件限制	293
7.2.1	外极几何限制	294
7.2.2	变换外极线为水平扫描线	295
7.2.3	线记号对应性的限制条件	296
7.3	求解对应性的匹配技术	299
7.3.1	基于灰度的匹配	300
7.3.2	松弛合作算法	301
7.3.3	概率松弛匹配技术	303
7.3.4	直线段匹配	305
7.3.5	曲线段匹配	311
7.4	双目视觉中的相机系统模型	312
7.4.1	基本概念	312
7.4.2	双目横模型结构	315
7.4.3	双目轴模型结构	318

7.5 小结	322
第八章 图像理解系统	
8.1 概述	324
8.1.1 引言	324
8.1.2 图像理解模型	325
8.2 面向区域的二维解释系统	327
8.2.1 航片解释系统	327
8.2.2 基于规则的景物解释系统	329
8.2.3 VISIONS	332
8.2.4 SCORPIUS	338
8.3 三维景物的解释系统	346
8.3.1 积木世界的理解	346
8.3.2 ACRONYM	350
8.3.3 三维 MOSAIC 系统	356
8.4 图像解释环境(IUE)	362
8.4.1 什么是图像解释环境	362
8.4.2 CME	366
8.4.3 KBVision	369
8.4.4 Power Vision	372
8.5 图像理解的系统结构	375
8.5.1 图像理解的不同层次和功能要求	377
8.5.2 图像理解与并行处理结构	378
8.5.3 典型的图像理解系统结构	382
8.5.4 图像理解结构的试验程序	386
8.6 小结	392
参考文献	394

第一章 绪 论

1.1 概 述

1.1.1 什么是图像理解

图像理解是研究用计算机系统解释图像,实现类似人类视觉系统理解外部世界的一门科学。所讨论的问题是,为完成某一任务需要从图像中获取哪些信息,以及如何利用这些信息获得必要的解释。图像理解的研究显然要涉及或包含研究获取图像的方法、装置和具体应用的实现,这就形成了所谓计算机视觉。计算机视觉是目前正在蓬勃发展的人工智能中的一个重要领域。

人类或其它高等生物的视觉过程,要解决的两个最基本的问题是,在视场中有什么及其在哪儿?这是一个复杂的信息处理过程,大体可分为如下两个部分:

感知 + **信息处理**

对生物来说,靠眼睛感知。不同种类的生物,在眼睛的个数和性能之间有很大的差异。一般高等生物都有二只眼睛,但也有多于二眼的生物:例如有三只眼的蜥蜴,有四只眼的四眼鱼,甚至还有八只眼睛的跳蜘蛛等。由于眼睛的个数、位置和瞳孔的形状及其变化,形成了各类生物独特的视觉机构,实现了千差万别的视觉功能。生物的眼睛虽然千差万别,但基本结构相同,都是由角膜、虹膜

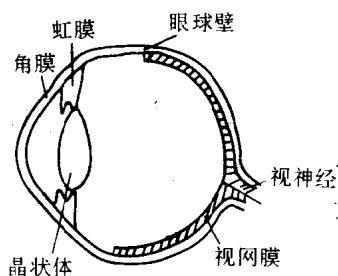


图 1.1.1 眼睛的基本结构



图 1.1.2 需要先验知识才能解释图像中景物的例子

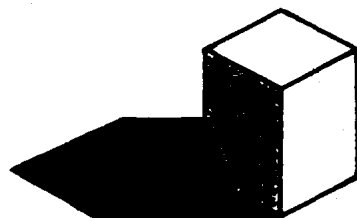
(构成瞳孔)、晶状体、视网膜、眼球壁和视神经组成。光通过瞳孔、晶状体在视网膜上被视细胞接收,并通过视神经通路传至大脑。视网膜及视觉通路对信息作了多层次处理,大脑有关部分则完成主要的信息处理任务。

计算机视觉也有相仿的过程:

成像 + **信息处理**



(a)



(b)

图 1.1.3 景物与成像
(a)投影成像; (b)原景物

成像用的装置可以是:光学照相机、红外照相机、X光、超声、电视摄像、综合孔径雷达成像、微波成像、电子显微镜、天文望远镜、热成像、同位素成像等。信息处理可粗略认为由图像理解系统担负。图像理解系统有几个鲜明的特点:①分阶段的信息处理带来了信息的多层表示;②对图像的解释,是以某种形式的描述实现的;③图像的正确解释离不开知识的导引。这最后一点,甚至对人的视觉理解也适合,即经验对视觉理解的重要性。现在举两个例子。请看图 1.1.2,如果事先不作任何说明,人

们很难想象出图中有一条黑白纹相间的花狗。再一个例子是图 1.1.3(a)所示的黑白二值图像,如果事先不知道光是从一侧面照射景物中目标时,就难于想象黑的部分为目标及其阴影所形成,即难

于理解图 1.1.3(b)的原景物图像。

常见的多层图像信息表示是低层,中层和高层描述,图 1.1.4 画出了三层图像理解的方块图。图中 A 处为输入的二维图像阵列, B 处为低层理解的结果,通常是以像

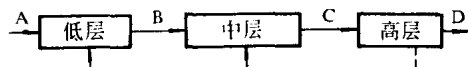


图 1.1.4 多层次的图像理解

元为单元测出的图像特征; C 处为中层理解的结果,它是低层描述的编组、抽象后形成的符号描述,减少了数据量,提高了描述的质量,更接近图像的本质。高层理解主要完成解释图像的任务,这里有以中层符号描述为基本单元的、反映景物与目标特性的模型,和服务于解释的知识库。 D 处为系统对输入图像的解释。人们常将这种由低层到中层直至高层输出的解释过程称为由底向顶的方式。自然也存在另一种方式:由上至下的方式,即先在 D 处根据模型对输入作一个假设的解释,以后去找证实此假设成立的中低层描述。如果将由底向顶和由上至下两种方式结合起来,即成混合的方式。例如可以先采用由底向顶的方式去解释,再将不满意的解释信息反馈到中低层,由此去调整中低层理解,最终实现满意的解释。到底采用什么方式,是设计一个图像理解系统必须要考虑的问题之一。图像理解的目的是解释一幅图像,这必然要涉及到识别图像中的目标,说明图像中各目标之间的关系、测量出各目标的相对尺寸和离观察点(如相机)的距离等。显而易见,目标的描述是其中一项主要内容。若仅仅考虑利用景物的二维投影时,利用区域(或边界)及其属性就可以描述目标。而若考虑三维情况,就要涉及三维目标的几何表示问题。由此而引来不同的图像理解的设想。这方面有代表性的是 Marr 的 $2\frac{1}{2}D$ ($2\frac{1}{2}$ 维)骨架、Barrow 和 Tenenbaum 的内在图像(intrinsic image)描述和 Lowe 的感知编组

识别理论。

Marr 的理论认为,信号可以分为四个层次:①图像,基元是像元的灰度值;②基本骨架,它反映了二维图像上明确的重要特征,如灰度变化的分布与组织,其基元是二阶导数过零点、块、终点与不连续点、边缘段……;③ $2\frac{1}{2}D$ 骨架,在以观察点为中心的坐标系统中,表示可见表面的外法线方向及离观察点的距离(深度),它们不连续点的轮廓等,其基元是局部表面外法线方向,离观察点的距离、深度不连续性及表面不连续性等;④三维模型表示,目的在以目标为中心的系统中描述三维目标的形状和它们在空间的组织,基元是体积的或表面的基元。图 1.1.5 表示了这个理论的主要部分。

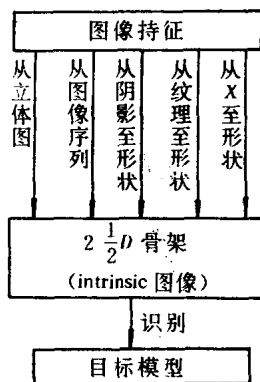


图 1.1.5 基于 Marr 计算理论的方块图

Barrow 和 Tennenbaum 提出用图像中可见表面的内在的表面性质来描述景物,这些性质诸如离观察点的距离,表面外法线方向、反射和照射等,构成了所谓内在的或实在的图像。图 1.1.6 表示了一个景物及其对应的内在图像。

Lowe 提出用人类感知的经验去描述目标。图 1.1.7 给出了他的主要设想。图中实线部分表示,图像特征由感知组织进行感知编组,而后形成相应的目标模型。解释的过程可以是由底向顶的,与模型匹配识别;也可以是由上至下提出假设,再在输入图像特征上找证据证实。这个过程无需三维目标表面的测量。然而由方块图中虚线部分可以看出,利用这些感知原理也可能推算出有关表面的信息。

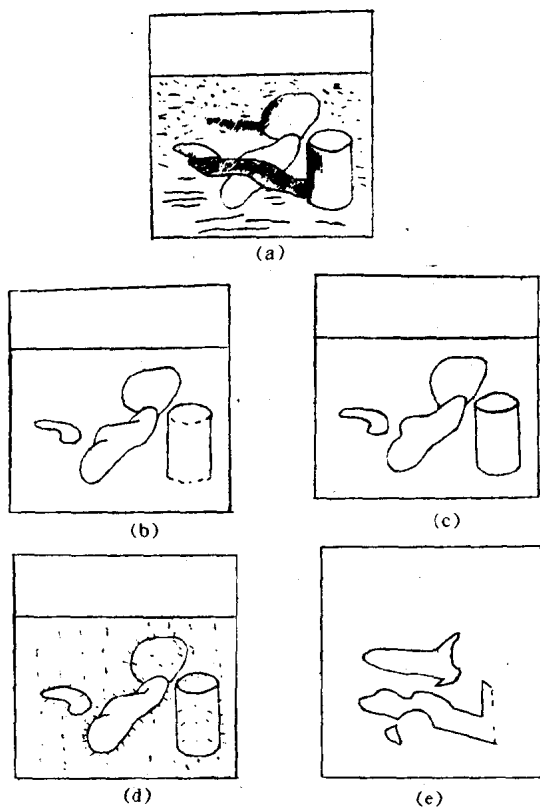


图 1.1.6 可见表面的内在图像

(a)原景物;(b)距离图像;(c)反射图射;(d)表面法线(矢量图);(e)照度图

由上可见,为实现图像理解,首先要研究系统构成的体制,这涉及到如何模型化目标、景物,如何作推断、解释;其次要研究实现

系统体制的各层符号描述,知识表示的方法和各自对应的算法及相互之间的转换。第三,要研究实现这些理论和算法的物理手段,即选用与研究有关的硬设备和系统等。

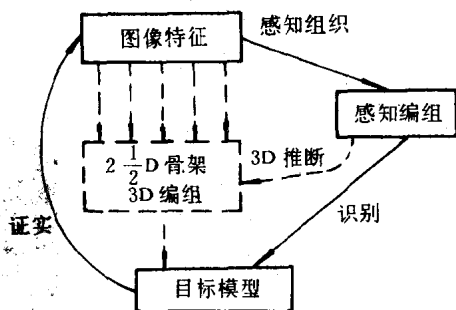


图 1.1.7 Lowe 的感知设想

1.1.2 图像理解的研究和应用

虽然对正常人来说,视觉理解外部世界并不是一件困难的事,然而要建立一个计算机系统,实现类比人视觉的计算机视觉,却是一件艰难的事。这里存在着以下几个主要困难:

(1)由于图像是三维景物的二维投影,一幅图像本身不具备复现三维景物的全部几何信息。图 1.1.8 表明,不同的空间四边形,即使人的单眼也难于区分。为克服这些困难,必须作合适的假定或附加新的

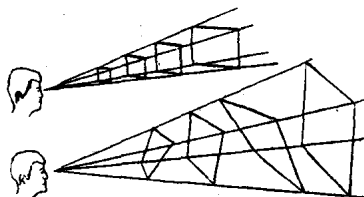
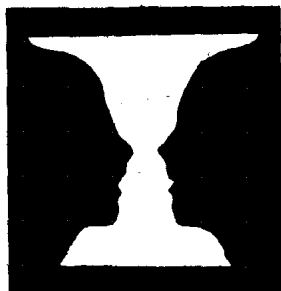


图 1.1.8 投影成像中的含糊性

测量(例如双目图像或多视点图像),这又同时引入新的计算问题。

(2)图像在形成过程中受到许多因素的影响,诸如照相机的质量,光传播介质的特性,照射景物光的特性、目标的反射特性等等,很难具体分清每个因素对某具体图像所产生的具体影响。

(3)图像本身不具有精确理解三维景物的全部信息,这就需要知识的导引。前面已有过些例子,图 1.1.9 是另一个例子。若没有知识的指引,我们很难说这是两个颠倒叠放的烛台型艺术品的剪影,还是一对好友深夜仍在窗前亲密对话的剪影。我们知道如何表示和应用所需要的知识并非一件易事,这正是人工智能中正致力解决的知识工程问题。



(4)计算机视觉是模仿人图 1.1.9 需要知识帮助解释的剪影的视觉,人的感知机理必然影响着计算机视觉的研究。例如,什么是感知的初始基元,基元是如何组织的,局部与全局感知的关系,优先敏感的结构、属性和时间特征等,这些都是心理学和神经心理学正着力研究的课题和领域。除此以外,计算机系统应用什么途径去“仿”,才能充分发挥自身的优点,这也是要下功夫探讨的问题。

(5)实现图像理解,在工程上面临可观的待处理的数据量。例如,对低分辨率的 128×128 黑白图像,要求约 16 千位的数据量;对高分辨率彩色 512×512 的图像,则要求约 750 千位;如果要处理 30 帧/秒的图像序列,则每秒要 500 千位~22.5 兆位。

可见图像理解是一项艰巨的任务,需要多学科长期的集体努力才能完成的。另一方面,技术的发展使得许多领域迫切需要应用图像理解,诸如完成各种工业或军事任务的机器人的视觉,进行资源勘探、天气预报、军事侦察的航空或卫星照片的自动分析,导弹的精确制导,战场的动态分析,天文学上改进星体图像及星体化学组成的研究,医学上器官异常的诊断及病理、细胞、染色体的分析,化学上分子结构组成分析,物理学上粒子轨迹及其研究等等。图像